RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10 516,317

Source: PCT

Date Processed by STIC: 1-17-06

ENTERED

PCT

```
        4 <110> APPLICANT: MIYAWAKI, Atsushi
        5         KARASAWA, Satoshi
        7 <120> TITLE OF INVENTION: CHROMOPROTEIN
        9 <130> FILE REFERENCE: P26359
C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/516,317
       12 <141> CURRENT FILING DATE: 2004-12-10
       14 <150> PRIOR APPLICATION NUMBER: JP2002/168583
       15 <151> PRIOR FILING DATE: 2002-06-10
       17 <150> PRIOR APPLICATION NUMBER: U.S. National Phase of PCT/JP03/07336
       18 <151> PRIOR FILING DATE: 2003-06-10
       20 <160> NUMBER OF SEQ ID NOS: 22
       22 <170> SOFTWARE: PatentIn version 3.3
       24 <210> SEQ ID NO: 1
       25 <211> LENGTH: 232
       26 <212> TYPE: PRT
       27 <213> ORGANISM: cnidopus japonicus
       29 <400> SEQUENCE: 1
       31 Met Ala Ser Lys Ile Ser Asp Asn Val Arg Ile Lys Leu Tyr Met Glu
       32 1               5                   10                  15
       35 Gly Thr Val Asn Asn His His Phe Met Cys Glu Ala Glu Gly Glu Gly
       36             20                  25                  30
       39 Lys Pro Tyr Glu Gly Thr Gln Met Glu Asn Ile Lys Val Thr Lys Gly
       40         35                  40                  45
       43 Gly Pro Leu Pro Phe Ser Phe Asp Ile Leu Thr Pro Asn Cys Gln Tyr
       44     50                  55                  60
       47 Gly Ser Val Ala Ile Thr Lys Tyr Thr Ser Gly Ile Pro Asp Tyr Phe
       48 65                  70                  75                  80
       51 Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Ile Tyr
       52                 85                  90                  95
       55 Glu Asp Gly Ala Tyr Leu Thr Thr Gln Gln Glu Thr Lys Leu Asp Gly
       56             100                 105                 110
       59 Asn Cys Leu Val Tyr Asn Ile Lys Ile Leu Gly Cys Asn Phe Pro Pro
       60         115                 120                 125
       63 Asn Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Cys Cys
       64     130                 135                 140
       67 Glu Met Arg Tyr Thr Arg Asp Gly Val Leu Cys Gly Gln Thr Leu Met
       68 145                 150                 155                 160
       71 Ala Leu Lys Cys Ala Asp Gly Asn His Leu Thr Cys His Leu Arg Thr
       72             165                 170                 175
       75 Thr Tyr Arg Ser Lys Lys Ala Ala Lys Ala Leu Gln Met Pro Pro Phe
       76             180                 185                 190
       79 His Phe Ser Asp His Arg Pro Glu Ile Val Lys Val Ser Glu Asn Gly
       80             195                 200                 205
```

Input Set : **A:\P26359.ST25.txt**
Output Set: **N:\CRF4\01172006\J516317.raw**

```
83 Thr Leu Phe Glu Gln His Glu Ser Ser Val Ala Arg Tyr Cys Gln Thr
84     210             215             220
87 Cys Pro Ser Lys Leu Gly His Asn
88 225             230
91 <210> SEQ ID NO: 2
92 <211> LENGTH: 699
93 <212> TYPE: DNA
94 <213> ORGANISM: cnidopus japonicus
96 <400> SEQUENCE: 2
97 atggcttcca aaatcagcga caatgtacgt atcaagttat atatggaggg cacagtcaac      60
99 aatcatcact tcatgtgcga agctgaagga gagggcaagc catacgaggg aactcaaatg     120
101 gagaacataa aagtcaccaa aggaggccct ctgccgttct cttttgatat cttgacgcct     180
103 aactgccaat atggaagcgt agccataacc aagtatacat cagggattcc agactacttt     240
105 aagcaatctt ttcctgaagg atttacctgg gaaagaacca caatctacga agatggggct     300
107 taccttacaa ctcaacaaga aaccaaactt gatggaaatt gcctcgtcta caatattaaa     360
109 atccttggat gtaattttcc ccccaatggt cctgtgatgc agaagaaaac ccaaggctgg     420
111 gaaccctgtt gcgagatgcg ctatacacgt gatggtgtgc tatgtggcca aacattaatg     480
113 gcacttaaat gcgccgatgg gaaccacctc acttgccatc tgagaactac ttacaggtcc     540
115 aaaaaggcag caaaggcgtt gcagatgcca cccttccatt tttcagacca tcgtcctgaa     600
117 atagtgaagg tttcagagaa cggcacacta tttgaacagc acgaaagttc agtggccagg     660
119 tactgtcaaa catgcccatc taaacttggt cacaattaa                           699
122 <210> SEQ ID NO: 3
123 <211> LENGTH: 17
124 <212> TYPE: DNA
125 <213> ORGANISM: Artificial Sequence
127 <220> FEATURE:
128 <223> OTHER INFORMATION: primer
131 <220> FEATURE:
132 <221> NAME/KEY: misc_feature
133 <222> LOCATION: (3)..(3)
134 <223> OTHER INFORMATION: n represents inosine
136 <220> FEATURE:
137 <221> NAME/KEY: misc_feature
138 <222> LOCATION: (5)..(5)
139 <223> OTHER INFORMATION: s represents c or g
141 <220> FEATURE:
142 <221> NAME/KEY: misc_feature
143 <222> LOCATION: (6)..(6)
144 <223> OTHER INFORMATION: n represents inosine
146 <220> FEATURE:
147 <221> NAME/KEY: misc_feature
148 <222> LOCATION: (9)..(9)
149 <223> OTHER INFORMATION: n represents inosine
151 <220> FEATURE:
152 <221> NAME/KEY: misc_feature
153 <222> LOCATION: (10)..(10)
154 <223> OTHER INFORMATION: h represents a, t, or c
156 <220> FEATURE:
157 <221> NAME/KEY: misc_feature
```

RAW SEQUENCE LISTING   DATE: 01/17/2006
PATENT APPLICATION: **US/10/516,317**   TIME: 13:26:23

Input Set : **A:\P26359.ST25.txt**
Output Set: **N:\CRF4\01172006\J516317.raw**

```
158 <222> LOCATION: (12)..(12)
159 <223> OTHER INFORMATION: n represents inosine
161 <220> FEATURE:
162 <221> NAME/KEY: misc_feature
163 <222> LOCATION: (13)..(13)
164 <223> OTHER INFORMATION: s represents c or g
166 <220> FEATURE:
167 <221> NAME/KEY: misc_feature
168 <222> LOCATION: (15)..(15)
169 <223> OTHER INFORMATION: n represents inosine
171 <400> SEQUENCE: 3
```
W--> 172 ggngsnccnh tnscntt   **17**
```
175 <210> SEQ ID NO: 4
176 <211> LENGTH: 44
177 <212> TYPE: DNA
178 <213> ORGANISM: Artificial Sequence
180 <220> FEATURE:
181 <223> OTHER INFORMATION: primer
183 <400> SEQUENCE: 4
184 aactggaaga attcgcggcc gcagaatttt tttttttttt tttt          44
187 <210> SEQ ID NO: 5
188 <211> LENGTH: 36
189 <212> TYPE: DNA
190 <213> ORGANISM: Artificial Sequence
192 <220> FEATURE:
193 <223> OTHER INFORMATION: primer
196 <220> FEATURE:
197 <221> NAME/KEY: misc_feature
198 <222> LOCATION: (24)..(25)
199 <223> OTHER INFORMATION: n represents inosine
201 <220> FEATURE:
202 <221> NAME/KEY: misc_feature
203 <222> LOCATION: (29)..(30)
204 <223> OTHER INFORMATION: n represents inosine
206 <220> FEATURE:
207 <221> NAME/KEY: misc_feature
208 <222> LOCATION: (34)..(35)
209 <223> OTHER INFORMATION: n represents inosine
211 <400> SEQUENCE: 5
```
W--> 212 ggccacgcgt cgactagtac gggnngggnn gggnng   **36**
```
215 <210> SEQ ID NO: 6
216 <211> LENGTH: 22
217 <212> TYPE: DNA
218 <213> ORGANISM: Artificial Sequence
220 <220> FEATURE:
221 <223> OTHER INFORMATION: primer
223 <400> SEQUENCE: 6
224 agacgaggca atttccatca ag                                   22
227 <210> SEQ ID NO: 7
```

```
228 <211> LENGTH: 20
229 <212> TYPE: DNA
230 <213> ORGANISM: Artificial Sequence
232 <220> FEATURE:
233 <223> OTHER INFORMATION: primer
235 <400> SEQUENCE: 7
236 ggccacgcgt cgactagtac                                        20
239 <210> SEQ ID NO: 8
240 <211> LENGTH: 24
241 <212> TYPE: DNA
242 <213> ORGANISM: Artificial Sequence
244 <220> FEATURE:
245 <223> OTHER INFORMATION: primer
247 <400> SEQUENCE: 8
248 ggctacgctt ccatattggc agtt                                   24
251 <210> SEQ ID NO: 9
252 <211> LENGTH: 30
253 <212> TYPE: DNA
254 <213> ORGANISM: Artificial Sequence
256 <220> FEATURE:
257 <223> OTHER INFORMATION: primer
259 <400> SEQUENCE: 9
260 cgggatccga ccatggcttc caaaatcagc                             30
263 <210> SEQ ID NO: 10
264 <211> LENGTH: 36
265 <212> TYPE: DNA
266 <213> ORGANISM: Artificial Sequence
268 <220> FEATURE:
269 <223> OTHER INFORMATION: primer
271 <400> SEQUENCE: 10
272 ccggaattct taattgtgac caagtttaga tgggca                      36
275 <210> SEQ ID NO: 11
276 <211> LENGTH: 232
277 <212> TYPE: PRT
278 <213> ORGANISM: cnidopus japonicus
280 <400> SEQUENCE: 11
282 Met Ala Ser Lys Ile Ser Asp Asn Val Arg Ile Lys Leu Tyr Met Glu
283 1               5                   10                  15
286 Gly Thr Val Asn Asn His His Phe Met Cys Glu Ala Glu Gly Glu Gly
287             20                  25                  30
290 Lys Pro Tyr Glu Gly Thr Gln Met Glu Asn Ile Lys Val Thr Lys Gly
291         35                  40                  45
294 Gly Pro Leu Pro Phe Ser Phe Asp Ile Leu Thr Pro Asn Cys Gln Leu
295     50                  55                  60
298 Gly Ser Val Ala Ile Thr Lys Tyr Thr Ser Gly Ile Pro Asp Tyr Phe
299 65                  70                  75                  80
302 Lys Gln Ser Phe Pro Glu Gly Phe Thr Trp Glu Arg Thr Thr Ile Tyr
303                 85                  90                  95
306 Glu Asp Gly Ala Tyr Leu Thr Thr Gln Gln Glu Thr Lys Leu Asp Gly
```

```
307                100                105                110
310 Asn Cys Leu Val Tyr Asn Ile Lys Ile Leu Gly Cys Asn Phe Pro Pro
311         115                120                125
314 Asn Gly Pro Val Met Gln Lys Lys Thr Gln Gly Trp Glu Pro Cys Cys
315     130                135                140
318 Glu Met Arg Tyr Thr Arg Asp Gly Val Leu Cys Gly Gln Thr Leu Met
319 145                150                155                160
322 Ala Leu Lys Cys Ala Asp Gly Asn His Leu Thr Cys His Leu Arg Thr
323         165                170                175
326 Thr Tyr Arg Ser Lys Lys Ala Ala Lys Ala Leu Gln Met Pro Pro Phe
327         180                185                190
330 His Phe Ser Asp His Arg Pro Glu Ile Val Lys Val Ser Glu Asn Gly
331         195                200                205
334 Thr Leu Phe Glu Gln His Glu Ser Ser Val Ala Arg Tyr Cys Gln Thr
335     210                215                220
338 Cys Pro Ser Lys Leu Gly His Asn
339 225                230
342 <210> SEQ ID NO: 12
343 <211> LENGTH: 699
344 <212> TYPE: DNA
345 <213> ORGANISM: cnidopus japonicus
347 <400> SEQUENCE: 12
348 atggcttcca aaatcagcga caatgtacgt atcaagttat atatggaggg cacagtcaac     60
350 aatcatcact tcatgtgcga agctgaagga gagggcaagc catacgaggg aactcaaatg    120
352 gagaacataa aagtcaccaa aggaggccct ctgccgttct cttttgatat cttgacgcct    180
354 aactgccaac ttggaagcgt agccataacc aagtatacat cagggattcc agactacttt    240
356 aagcaatctt ttcctgaagg atttacctgg gaaagaacca caatctacga agatggggct    300
358 taccttacaa ctcaacaaga aaccaaactt gatggaaatt gcctcgtcta caatattaaa    360
360 atccttggat gtaattttcc ccccaatggt cctgtgatgc agaagaaaac ccaaggctgg    420
362 gaaccctgtt gcgagatgcg ctatacacgt gatggtgtgc tatgtggcca aacattaatg    480
364 gcacttaaat gcgccgatgg gaaccacctc acttgccatc tgagaactac ttacaggtcc    540
366 aaaaaggcag caaaggcgtt gcagatgcca cccttccatt tttcagacca tcgtcctgaa    600
368 atagtgaagg tttcagagaa cggcacacta tttgaacagc acgaaagttc agtggccagg    660
370 tactgtcaaa catgcccatc taaacttggt cacaattaa                          699
373 <210> SEQ ID NO: 13
374 <211> LENGTH: 232
375 <212> TYPE: PRT
376 <213> ORGANISM: cnidopus japonicus
378 <400> SEQUENCE: 13
380 Met Ala Ser Lys Ile Ser Asp Asn Val Arg Ile Lys Leu Tyr Met Glu
381 1               5                10                15
384 Gly Thr Val Asn Asn His His Phe Met Cys Glu Ala Glu Gly Glu Gly
385             20                25                30
388 Lys Pro Tyr Glu Gly Thr Gln Met Glu Asn Ile Lys Val Thr Lys Gly
389         35                40                45
392 Gly Pro Leu Pro Phe Ser Phe Asp Ile Leu Thr Pro Asn Cys Gln Met
393     50                55                60
396 Gly Ser Val Ala Ile Thr Lys Tyr Thr Ser Gly Ile Pro Asp Tyr Phe
397 65                70                75                80
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 01/17/2006
PATENT APPLICATION: US/10/516,317   TIME: 13:26:24

Input Set : A:\P26359.ST25.txt
Output Set: N:\CRF4\01172006\J516317.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; N Pos. 3,6,9,12,15
Seq#:5; N Pos. 24,25,29,30,34,35

VERIFICATION SUMMARY DATE: 01/17/2006
PATENT APPLICATION: **US/10/516,317** TIME: 13:26:24

Input Set : **A:\P26359.ST25.txt**
Output Set: **N:\CRF4\01172006\J516317.raw**

L:11 M:270 C: Current Application Number differs, Replaced Current Application Number
L:172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0
L:212 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0